DOCUMENT RESUME

ED 051 975 RE 003 662

AUTHOR Froese, Victor
TITLE Cloze Readability versus the Dale-Chall Formula.
PUB DATE Apr 71
NOTE 12p.; Paper presented at the meeting of the International Reading Association, Atlantic City, N.J., Apr. 19-23, 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Cloze Procedure, Comparative Analysis, Content Reading, Correlation, Grade 6, *Readability, *Reading Level, Reading Materials, *Science Materials, *Textbook Evaluation, Validity

ABSTRACT

The validity of the Dale-Chall Readability ratings for sixth-grade science textbooks when compared to an independent criterion of language difficulty expressed in cloze units was examined. The Canadian Lorge-Thorndike IQ Test and the Canadian Test of Basic Skills, Vocabulary and Comprehension subtests, were administered to 366 sixth graders. Mean IQ score was 102, while mean scores for vocabulary and comprehension were 6.19 and 6.15, respectively. Passages from 12 science textbooks were subjected to the cloze technique and administered to the subjects. Findings revealed that (1) the 12 cloze passages yielded a reliability coefficient of .79; (2) cloze scores correlated from .55 to .85 with the intelligence measure; (3) cloze scores correlated from .64 to .86 with vocabulary, and from .52 to .85 with comprehension; (4) the cross-validation coefficient of .90 obtained by Dale and Chall and the cross-validation coefficient of -.29 obtained in this study are significantly different beyond the .01 level of confidence. Major conclusions were (1) that the cloze tests are reliable measures of language difficulty and (2) that the Dale-Chall Readability Formula is not a valid measure of sixth-grade science textbook materials when the cloze procedure is used as a criterion. References are included. (VJ)

ED051975

Victor Froese
University of Manitoba
Winnipeg 19, Manitoba
Canada

U. S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

CLOZE READABILITY VERSUS THE DALE-CHALL FORMULA

Research Report
Thursday, April 22, 1971
2:15-3:15 P.M.

Need for study

Much contradictory evidence is found in the literature with regards to the measurement of readability. Some of the most pertinent findings with regards to the readability of elementary science textbook materials will be examined in this paper.

Since by generally accepted standards a readability formula is applicable only to material similar to the criterion on which it was based (3)(10), it appears that the Dale-Chall formula scores for elementary science textbook material should be used cautiously. Chall, in 1958, stated that no studies based

RE003 662

exclusively on the specialized subject matters of science or mathematics had been reported. A search of the literature to date revealed no validation studies done on elementary science materials.

However, several studies reported findings based on the application of the Dale-Chall formula to materials which were not included in the original validation or subsequent validations. This would lead one to question the results of these studies.

Brown (2) questioned the appropriateness of the Dale list of 3000 familiar words as a vocabulary load factor in the Dale-Chall formula. He found that students at the seventh and eighth grade levels apparently comprehended a N.S.T.A. publication entitled Spacecraft which was rated at 11-12 grade-level by the Dale-Chall formula. Furthermore, when the vocabulary from a 1961 edition of a third-grade science textbook was accepted as familiar, it was found that it lowered the readability of the same sample from 11-12 grade-level to 9-10 grade-level. His conclusion was that when the Dale-Chall formula was applied to science textbooks it seemed to place them higher than was warranted.

Walker (10) used the Dale-Chall formula to evaluate thirty-nine commercially produced programmed textbooks for grades four to six and found that the Dale-Chall readability levels for these books were consistently higher than the grade levels assigned to them by the publisher. Sixty-seven percent of the books were rated above the grade level indicated by the publisher and three of the books had no samples

at the intermediate level. This type of finding may be questioned since the variability may be due to either the nature of the formula or the material being evaluated. Although high relationships between the Fry, Lorge, Flesch, and Dale-Chall formulas have been reported by Fry (4), contradictory evidence is quoted by Michaelis and Tyler (13). Marshall (12), furthermore, found no significant relationship between readability and comprehension when using the Flesch Reading Ease formula on high school physics books.

Since most of the formulas are based on common or highly interrelated factors, and since the vocabulary factor accounts for a very substantial amount of the variance of readability elements, this aspect of the formulas must be examined for its contribution toward formula consistency or inconsistency. There is also some evidence to indicate that vocabulary difficulty is a better predictor at lower levels of difficulty (3) and for poorer readers (5).

The Winnetka, Dale-Tyler, Gray-Leary, Lorge, Dale-Chall, Dolch and Spache formulas all use basic word lists which are dated since they were compiled in the 1920's and 1930's. Evidence gathered by Stone (15) and Jacobs (7) indicates that vocabularies have changed. Kucera and Francis (6) also found that of the Dolch 220 words of the 1920's, 82 words (37%) were not among their 1960 compilation of the most frquently used 220 words. Recognizing the possibility that this change might affect the readability, Spache has now adopted Stone's revised list. The validity of the other formulas, however, could be questioned since their lists have not been revised.

Bormuth, in investigating a measure of word depth as a predictor of comprehension difficulty in literature, science, and history materials, found it to predict differences in comprehension of different subject matter written at the same readability level. This led him to suggest:

> In the past the assumption has been that such differences were caused by differences inherent in the content of the subject matters themselves. Though this concept nas never been rigorously defined, it is given the labels of concept difficulty or idea density. (1)

In other words he is questioning differences in readability commonly attributed to the content areas.

Chall, in her 1958 summary of readability research, stated the need for cross-validation research into textbook materials. She wrote:

> Most of the existing validation studies have been on juvenile fiction. Since the formulas are used by educational publishers and textbook committees for evaluating textbooks for a particular grade, we need to know how valid the predicted grade-placement indexes of the various formulas are when compared to tested comprehension or to teacher or pupil judgement of difficulty.(3).

In his 1963 investigation of readability research, Klare proposed the following areas for future research which are related to this study: The use of new word-lists, specialized word-lists, specialized lists for specialized purposes, longer lists, criterion passage refinement, when or where to use specific formulas, and separate norms for the readability of different types of materials and audiences (8).

It is apparent that the Dale-Chall formula has not been validated on elementary science materials. The Dale List of

3000 familiar words may be questioned for relevance to science material. And the weighting of the vocabulary factor in the formula may need adjustment for certain types of materials or for certain levels of materials.

Furthermore, differences in readability due to content materials have been questioned. Also some contradictory evidence about the relationship between readability and comprehension has been presented.

Finally, a need for readability research with respect to validation, use of word lists, and use on content material has been expressed.

Further research, especially on the validity of using the Dale-Chall formula on elementary science textbook material, is therefore needed.

The problem

This study was designed to investigate the problem: How valid are the Dale-Chall Readability ratings for sixth grade science textbook materials when compared to an independent criterion of language difficulty expressed in cloze units.

Null hypothesis

There is no significant difference between the correlation coefficient of .90 obtained in the original cross-validation of the Dale-Chall Readability Formula and the coefficient of correlation between the Dale-Chall readability ratings and the mean cloze score obtained by the subjects completing the cloze test over randomly selected sixth grade science textbook passages.

Definition of terms

1. CROSS-VALIDATION. A comparison of a predicted readability level to an independent criterion of readability (or language difficulty).
2. CLOZE PROCEDURE (or CLOZE TECHNIQUE). A method of intercepting a message from a transmitter, mutilating its language patterns by deleting parts [in this case, every 5th word] and so administering it to receivers that their attempts to make the pattern whole again yields a number of cloze units (the score) (16).
3. INDEPENDENT CRITERION of LANGUAGE DIFFICULTY. A measure of the language difficulty of a passage expressed in mean cloze units for the subjects having read a particular passage.
4. READABILITY. A prediction of the ease or difficulty with which written material may be comprehended according to the specified criteria.
5. SIXTH GRADE SCIENCE TEXTBOOK MATERIALS. Randomly selected samples of science materials from textbook series designated for the sixth grade by the publisher and listed in Textbooks in Print 1968 (17).

The population

The sample under study consisted of 366 randomly selected sixth grade students enrolled in thirteen classrooms in eleven different schools in School District #33, Chilliwack, British Columbia, Canada. The 366 students represented approximately one-half of the sixth grade students in this district.

The mean Canadian Lorge-Thorndike IQ (11) for the 366 students was 102 with a standard deviation of 15.

Approximately 51 percent were girls and approximately 49 percent were boys.

The Canadian Test of Basic Skills (9), subtest Vocabulary and Comprehension were administered sixty days before this study (in October) and resulted in a mean score of 6.19 (SD = 1.27) and 6.15 (SD = 1.26) respectively.

Both the IQ test and the Basic Skills test were administered as a regular part of the District testing program.

Summary of findings

Four specific findings were obtained from the data in this study:

1. The twelve cloze passages used in this study yielded a mean reliability coefficient of .79 with a range of .64 to .89. These coefficients are comparable to some standardized group reading tests.

2. The cloze scores obtained for each passage correlated from .55 to .85 with intelligence as measured by the Canadian Lorge-Thorndike test. These coefficients are comparable to those found in studies by Hafner and Jenkinson (14).

3. The cloze scores obtained for each passage correlated from .64 to .86 with reading vocabulary, and from .52 to .85 with reading comprehension. These findings compare favorably with those cited by Rankin (14).

4. The cross-validation coefficient of .90 obtained by Dale and Chall and the cross-validation coefficient of -.29 obtained in this study are significantly different beyond the .01 level of confidence.

Limitations

The findings and conclusions in this study must be interpreted keeping in mind the relatively small number of passages, the reliability and validity of the cloze tests, the specific samples of subjects, and the science material used:

1. An item sampling error was probably inherent in the cloze passages.

2. Approximately 55 percent of the passages were within the frustration levels of the examinees completing the passages. This factor could cause a leveling-off of the distribution of scores for the cloze passages at the upper end of difficulty.

Conclusions

Subject to the limitations of this study the following four conclusions were drawn from the findings:

1. The cloze tests are reliable measures of language difficulty when used for group testing.

2. The cloze tests are valid measures of language difficulty as demonstrated by their concurrent validity with intelligence test scores.

3. The cloze tests are valid measures of language difficulty as demonstrated by their congruent validity with reading vocabulary and reading comprehension scores.

4. The Dale-Chall Readability Formula is not a valid measure of sixth-grade science textbook materials when the cloze procedure is used as a criterion.

TABLE OF DATA FOR TWELVE RANDOMLY SELECTED ELEMENTARY SCIENCE TEXTBOOK PASSAGES

Passage Number	Mean Cloze Score	Cloze Standard Deviation	Corrected Split-half Reliability	Dale-Chall Readability Index	Correlation of Cloze & Intelligence	Correlation of Cloze & Vocabulary	Correlation of Cloze & Comprehension
1	16.45	6.17	.82	7.06	.67	.74	.64
2	15.78	5.04	.73	6.90	.85	.86	.85
3	18.40	4.78	.67	7.45	.55	.78	.79
4	12.30	5.75	.64	6.82	.62	.65	.66
5	18.08	6.45	.83	5.17	.67	.64	.58
6	19.90	5.11	.89	6.92	.66	.67	.71
7	17.10	6.01	.82	6.91.	.75	.77	.76
8	11.65	4.69	.80	7.70	.63	.64	.63
9	20.48	6.35	.82	8.01	.65	.79	.73
10	13.27	3.73	.75	8.38	.61	.64	.52
11	19.00	5.71	.73	5.54	.76	.82	.79
12	15.82	4.24	.64	6.33	.66	.83	.79

Implications

Since the null hypothesis was rejected, and since the Dale-Chall Readability Formula was found not to be a valid measure of sixth-grade science material, the following implications should be considered:

1. Previous research using the Dale-Chall Readability Formula on elementary science textbook materials should be re-examined.

2. In order to validly measure science materials with the Dale-Chall Readability Formula the following alterations may be necessary: (a) Weighting the formula differently, (b) compiling a specialized word list, (c) up-dating the word list, (d) developing special norms, or (e) some combination of these.

3. For some purposes, such as matching a book to a reader, the cloze procedure may be a more robust and more parsimonious measure of language difficulty than the Dale-Chall Readability Formula.

4. Reasearch using passages written specifically for a certain Dale-Chall readability level should be interpreted cautiously since this has, in effect, eliminated the readability measurement error. This could lead to different conclusions from those indicated when randomly selected passages are subjected to a readability formula.

5. This study should be replicated at the same and at other grade levels, with differing populations, and with other randomly selected science textbook passages in order that the results could be more generalizable.

References

1. Bormuth, J. R. "Mean Word Depth as a Predictor of Comprehension Difficulty," California Journal of Educational Research, XV (November 1964), 226-231.

2. Brown, W. R. "Science Textbook Selection and the Dale-Chall Formula," School Science and Mathematics, LXV (February 1965), 164-167.

3. Chall, J. S. Readability: An Appraisal of Research and Application, Bureau of Educational Research Monographs, No. 34 (Columbus, Ohio, 1958), p. 35, p. 45, p. 165.

4. Fry, E. "A Readability Formula That Saves Time," The Journal of Reading, XI (April 1968), 513-516.

5. Holmes, J. A. and Singer, H. Speed and Power of Reading in High School, Bureau of Educational Research and Development (U. S. Government Printing Office, Washington, 1966), pp. 104-109.

6. Johnson, D. D. "The Dolch List Reexamined," The Reading Teacher, XXIV (February 1971), 449-457.

7. Jacobs, H. D. "A Replicative Investigation of the Buchingham-Dolch Free-Association Word Study," U. S. O. E. Project 7-8037 (Doctoral dissertaion, University of Oregon, 1967).

8. Klare, G. R. The Measurement of Readability, (Ames, Iowa: Iowa State University Press, 1963), pp. 189-190.

9. King, E. M., Lindquist, E. F. and Heironymous, A. N. Canadian Test of Basic Skills, Test V: Vocabulary and Test R: Reading Comprehension (Toronto, Ontario: Thomas Nelson & Sons, 1968).

10. Lorge, Irving "Readability Formulae--An Evaluation," Elementary English, XXVI (February 1949), 87-88.

11. Lorge, I., Thorndike, R. L., Hagen, E. and Wright, E. N. Canadian Lorge-Thorndike Intelligence Tests, Level D, Verbal, (Toronto, Ontario: Thomas Nelson & Sons, 1968).

12. Marshall, J. S. "Comprehension and Alleged Readability of High School Physics Textbooks," Science Education, XLVI (October 1962), 335-346.

13. Michaelis, J. U. and Tyler, F. T. "A Comparison of Reading Ability and Readability," Journal of Educational Psychology, XLII (1951), 491-498.

14. Rankin, E. F. "The Cloze Procedure--A Survey of Research," The Philosophical and Sociological Bases of Reading, Ed. E. L. Thurston and L. E. Hafner, Fourteenth Yearbook of the National Reading Conference (Milwaukee, Wisconsin: 1965), pp. 133-150.

15. Stone, C. R. "Measuring Difficulty of Primary Reading Material: A Constructive Criticism of Spache's Measure," Elementary School Journal, LVII (October 1956), 36-41.

16. Taylor, W. L. "Cloze Procedure: A New Tool for Measuring Readability," Journalism Quarterly, XXX (1953), 416.

17. Textbooks in Print/1968, (New York: R. R. Bowker Co., 1968), pp. 217-250.

18. Walker, W. L. "Measured Readability of Intermediate Grade Textbooks," The Teachers College Journal, XXXVII (March 1966), 179-181.